

BEFORE THE
POSTAL REGULATORY COMMISSION
WASHINGTON, DC 20268-0001

Inquiry Concerning Service)	
Performance Measurement Data)	Docket No. PI2016-1

**REPLY COMMENTS OF
JOHN HALDI, PH.D.**
(February 9, 2016)

Comments in response to Order No. 2791 were submitted on December 14, 2015 by the Public Representative (“PR”), and joint comments were submitted by the Association for Postal Commerce (“PostCom”) and Major Mailers Association (“MMA”). Each is discussed in turn.

I. PR Comments

A. PAEA Mandates Reporting on “Reliability,” but What That Means for the Performance Measurement and Reporting System Is Subject to Different Interpretations.

The PR’s Comments note that “this docket is an outgrowth of a recent report by the United States Government Accountability Office (GAO).”¹ *Id.* at 1. The PR then cites the regulatory framework for performance measurement, including the following statement:

The statutory basis for service performance measurement and reporting is contained in three sections of title 39 of the United States Code: sections 3691, 3652(e), and 3653. 39 U.S.C. § 3691, § 3652, and § 3653. ... Section 3652(a)(2)(b)(i) requires the Postal Service to include in its Annual Compliance Report

¹ Actions Needed to Make Delivery Performance Information More Complete, Useful, and Transparent. U.S. Gov. Accountability Office, GAO-15-756, U.S. Postal Service: Actions Needed to Make Delivery Performance Information More complete, Useful, and Transparent. (2015) (GAO Report).

(ACR) measures of the quality of market dominant produce [sic] service, “including...the level of service (described in terms of speed of delivery **and reliability**)....” In that connection, section 3652(e)(1) gives the Commission the authority to prescribe the content and form of the Postal Service’s reports, which includes the Postal Service’s ACR. [*Id.* at 2 (emphasis added).]

The statutory requirement for **reliability**, cited by the PR, raises important issues.² The term “reliability” is separate and distinct from “speed of delivery.” Presumably, therefore, it refers to something other than speed of delivery, which is measured by timeliness of delivery against an established target. However, the statute neither defines nor explains what is intended by the term “reliability” as used therein. In the context of (i) actual delivery services provided by the Postal Service, and (ii) performance measurement and reporting thereof, the interpretation of “reliability” can be subject to different interpretations. Succinctly, reliability can refer to:

1. *Accuracy and completeness of data* recorded in the performance measurement system;³ or
2. *Predictability and consistency of the actual mail delivery service* provided by the Postal Service for the various categories of mail;
or

² The PR also discusses the issue of whether collected data are *representative* of each specified category. Representativeness is a fundamental threshold issue of critical importance, as the PR recognizes. In prior years some mailers have occasionally submitted data depicting delivery service that was extremely erratic and inconsistent. The Postal Service inevitably would argue that such non-Postal Service data should be disregarded completely because they were *not representative*. With the Postal Service now responsible for design and implementation of the data collection system, representativeness, although still critically important, hopefully will cease to be an issue.

³ As data for a category of mail become more *complete*, the data generally are deemed to be more reliable, especially if the data can be considered a “census” rather than a “sample.”

3. Both of the above, meaning information (*e.g.*, charts, tables, analyses) designed to focus on and depict whether actual performance for each category of mail has been less than reliable as a means of driving operational improvements as and where needed.

Differences between the preceding descriptions of reliability are significant. By way of illustration, for any category of mail and with sufficient resources it theoretically is possible to measure and record with a high degree of *accuracy and completeness* data on the actual performance of delivery service that itself may be *highly erratic, inconsistent, unpredictable* and, from the viewpoint of mailers, extremely *unreliable*. That is, the data may be accurate, hence reliable (in the sense of data quality), while the delivery service itself is somewhat unreliable. And by the same token, although the data may be 100 percent correct, the performance reporting system may fail to provide Congress, mailers and other interested parties with charts, tables or other information that demonstrate — accurately and irrefutably — just how inconsistent and unreliable performance is for each category of mail.⁴

⁴ This routine of “reliable data/unreliable service” is not intended as some semantical play on words similar to the “Who’s on first” routine made famous by comedians Abbott and Costello. The Commission’s Notice appears focused almost entirely on obtaining data that are accurate and more comprehensive. In that respect, the Commission can be said to be doing a good job. It is important to note, though, that if the statute — and Congressional interest — pertains to *reliability of delivery service*, then to date the Commission’s focus on data quality misses the main point. The Commission neither has developed on its own initiative, nor required from the Postal Service, any submission depicting just how unreliable, erratic, or inconsistent performance has been for any category of mail. *See* discussion of the GAO Report, *infra*. In the service performance measurement and reporting system, measures that explicitly depict unreliability for any area or category of mail are scant and wanting. Goals for reducing the extent of unreliability, *i.e.*, for improving reliability of delivery service, are likewise scant.

In other words, the reporting format may fail to elucidate, even obfuscate, the extent of *inconsistent, unreliable service* provided by the Postal Service. Should that occur, then no matter how accurate and complete are the data, the value of the reporting system is diminished. Even if existing tables of performance data are expanded to the point where the data therein become 100 percent complete (*i.e.*, a census, not just a large sample), and hence unarguably *representative*, they still might fail **to diagnose and highlight** the issue of how just unreliable is the actual delivery service which the Postal Service provides to any area or category or mail.⁵

When mail is delivered on-time, vis-a-vis established service standards, neither speed of delivery nor reliability are issues. Inconsistency and unreliability of service (the “flip side” of reliability) pertain to mail *not* delivered on-time, *i.e.*, to the tail-of-the-mail. To date, the Postal Service’s service performance measurement and reporting system focuses chiefly on speed of mail service, as measured by on-time delivery. Reliability as measured by accepted statistical measures of inconsistency, such as (i) average days late for mail not delivered on-time, or (ii) the standard deviation of days late for that portion of mail not delivered on-time is virtually nonexistent. Inasmuch as measures of reliability are not presented, **targets for reliability do not exist.**

Note that identifying unreliability of delivery to any area (*e.g.*, rural areas, sparsely populated states) or for any category of mail (*e.g.*, First-Class letters, Standard Mail, or

⁵ This is not to say that the Commission’s efforts to expand the quantity and improve quality of performance data constitute a fool’s errand. Rather, when “perfection” of the data finally is achieved, the task of using the data to identify systemic pockets of unreliability and making performance data more operationally meaningful still will be far from complete.

Periodicals) is a first step toward service improvement.⁶ The Postal Service then can use such information to improve consistency of its delivery service. Accurately recording performance data clearly is of fundamental importance. However, it should not be expected that accurate data, *per se*, will generate automatically the information required for insightful reports on unreliability.⁷

By way of brief summary, a *narrow view* of the statutory requirement for reliability focuses on and *is limited to* having good data, which are *necessary* to drive service improvement. A *broad view* of the statutory requirement for reliability recognizes that data alone are *not sufficient*, and explicit statistical measures of reliability are needed.

B. The Interpretation and Definition of Reliability by the PR (and the Commission) Seems Focused Exclusively on Data Quality, Not Reliability of Delivery Service

The PR does not offer an explicit interpretation of what the statute intends by its reference to “reliability.” Instead, citing to texts on statistics, the PR offers the following definition of *reliability* (at 9-10).

Reliability is another very important element of **testing data quality**, and it is usually defined as reproducibility and stability

⁶ Organizations typically find it difficult to solve problems that have not been identified and diagnosed.

⁷ Performance measurement by the Postal Service can be viewed as still in its formative stage. The Postal Service has been delivering mail for well over 200 years. Throughout most of those years, after mail arrived at the local post office it was delivered to addressees, which was deemed sufficient. Prior to 2007, the Postal Service made little effort to measure performance in terms of either timeliness or consistency. Except for Single-piece First-Class Mail and Express Mail, essentially no effort was made. As the PR’s comments point out, only in 2007 did the Postal Service issue “regulations establishing service standards for market dominant products. This was the first step in adopting a system of service performance measurement and reporting.” *Id.* at 3.

(consistency) of the obtained **measurement estimates** and/or scores. [*Id.* at 9 (emphasis added).]

Regrettably, this is a statistical definition applicable only to data quality. It has little to do with **reliability of delivery service** provided by the Postal Service. This definition makes no mention of how the unreliability of delivery service being provided to any category of mail either should be, or might be, measured. It would appear that the PR's concerns about reliability are limited solely to accuracy and completeness of the data (issues raised in the Commission's Notice), not to whether the reporting format is capable of demonstrating the extent to which any category of service provided by the Postal Service can be considered unreliable and inconsistent. Citing the Commission's Notice in Order No. 2791, the PR seemingly endorses a somewhat *narrow view* of reliability, in line with the above definition:

the Commission identified three principle subjects on which it seeks information regarding the quality and completeness of **service performance data** provided by the Postal Service: (1) potential deficiencies in the accuracy, **reliability**, and representativeness of **service performance measurement data**. ...” [*Id.* at 7 (emphasis added).]

The PR also fails to discuss that the delivery service itself may be in flux, with delivery becoming more or less consistent over time. Should that be the case, one should not expect performance data to be more stable or consistent than the underlying actual delivery service being measured. In other words, performance data should reflect the underlying reality as it changes, not necessarily be expected to reflect reproducibility of the measurement estimates, as indicated by the PR's definition.

C. The Annual Compliance Report Should Include Measures of the Reliability of Delivery Service Provided to Market Dominant Products

The PR cites the statute's requirement to include in the Annual Compliance Report (ACR):

measures of the quality of market dominant produce [sic] service, "including...the level of service (described in terms of speed of delivery **and reliability**). [PR Comments at 2 (emphasis added).]

Surely this language in the statute intends for the ACR to include information on the reliability of delivery service, not a discourse on statistical quality of performance data.

Admittedly, the 2015 ACR was not available to the PR at the time his Initial Comments were drafted, but reference to the 2014 ACR would have served equally well. As explained herein, each ACR to date regularly avoids any mention of the possibility that delivery service may not always be as reliable as desired.

1. Measures of Service Performance in the FY 2015 ACR are Inadequate and Fail to Comply with the Statute

With respect to service performance, this year's ACR states that

During FY 2010, the Commission issued its final rules on periodic reporting of service performance measurement and customer satisfaction, which are codified at 39 C.F.R. Part 3055.28 Among other things, Commission Rules 3055.20 through 3055.24 require annual reporting of service performance achievements at the national level for all market dominant products. Reporting, however, is not required where the Commission has granted a semi-permanent exception or a temporary waiver. **The Postal Service's report, including information responsive to the criteria listed in Rule 3055.2(b)-(k), is included as USPS-FY15-29.** [FY 2015 ACR at 56 (emphasis added).]

USPS-FY-15-29, Service Performance of Market Dominant Products, states that “USPS-FY15-29 contains the annual report on the service performance of market dominant products.” It lists “five Microsoft Excel spreadsheets with **data required by Part 3055, Subpart A, of the Commission’s Rules.**” (Emphasis added.) These spreadsheets contain all the data which the Postal Service reports in the ACR for service performance. The data in those spreadsheets purportedly comply with and satisfy Part 3055, Subpart A, of the Commission’s Rules.

2. Data in the Postal Service’s Service Performance Spreadsheets Accompanying the FY 2015 ACR Pertain Only to Speed of Delivery, as Measured by On-time Delivery

For Standard Mail, the spreadsheet shows for each product the national target for on-time delivery (91 percent), and the actual achievement versus the established target (the annual on-time performance ranged from a low of 75.1 percent to a high of 98.1 percent). On-time delivery essentially measures speed of delivery as required by the statute, and the spreadsheets contained in USPS-FY-15-29 satisfy that part of the statute.

3. The FY 2015 ACR Contains No Statistical Measures for Reliability of Delivery Service, No Targets for Reliability of Delivery Service, and No Data Indicative of Actual Reliability of Delivery Service

In the FY 2015 ACR there exists a *total void* as regards reporting on reliability of delivery service, despite the statute’s requirement for:

the Postal Service to include in its Annual Compliance Report (ACR) measures of the quality of market dominant product service, “including...the level of service (described in terms of ... speed of delivery **and reliability**)....”

It is not clear whether the shortcoming is in the Commission's rules in Part 3055, Subpart A, or the Postal Service's reporting. Regardless, this year's ACR contains nothing whatsoever about reliability of delivery service — no targets, no statistical measures, and no data. This failure to comply with the statute should be rectified by the time the FY 2016 ACR is submitted.

D. Reporting of Performance Data Could Be More Transparent

Performance data collected by the Postal Service and provided in tables available on the Commission's web site already are rather extensive. The GAO Report indicates that about 55 percent of the volume of all market dominant products is now in measurement. This means the Postal Service annually records observations on tens of millions of pieces of mail. A plethora of data already exist. And it continues to grow each year. More interpretation of the data would be helpful, yet "the Postal Service does not provide any explanations of the differences in the margins of error, and/or the reliability of the reported scores." as the PR observes. *Id.* at 9. Here the PR is on to something.

The lack of transparency is consistent with a *narrow view* of the statutory requirement for reliability, which makes data quality the centerpiece of the discussion. Lack of transparency constitutes a gap in current design of the service performance and reporting system, as the GAO Report points out (see below). The limited scope of the present docket appears to preclude comments or suggestions for development of better reporting formats whose aim would be to use existing data to highlight unreliable delivery service. Consequently, the Commission should consider opening a new docket pertaining to improved formats for reporting of service performance.

E. The GAO Report Offers a Commendable Illustration of One Way to Analyze Unreliability

The GAO Report (*see* fn. 1 *supra*) obviously adopts a *broad view* of the statutory requirement for reliability, *i.e.*, reliability of actual delivery performance, not just reliability of data. The GAO Report uses existing data to construct a chart that illustrates the range of variation in quarterly performance for one category of mail; namely, Single-Piece First-Class Letters and Postcards with a 3-to-5 day Delivery Standard.⁸ For that category of First-Class Mail, and in the second quarter of FY 2015, on-time delivery ranged from a low of 44 percent to a high of 80 percent — hardly a consistent and reliable level of service. Among 64 postal districts studied, even the highest score was well below the 95 percent target. *Id.* at 24.

Neither the results nor details pertaining to conduct of that GAO study are important here. What's important is that the GAO Report illustrates one possible way of mining existing performance data to illustrate the extent of unreliability that mailers can face. Other, still better ways to illustrate the extent of unreliable delivery service may well exist. *None, however, are contained in the performance measurement and reporting system as it now exists.* Nor do any seem to be under development. The performance data base already represents a considerable effort by both the Postal Service and the Commission. The next step, using those performance data to diagnose, routinely and systematically, issues and areas in need of improvement, will require still more hard work, as the GAO Report illustrates.

Consistent with the *broad view* of reliability adopted by the GAO, their Report recommends that the Commission take actions to make delivery performance more complete,

⁸ *See* Report, *Op. Cit.*, Fig. 5 at 25.

useful, and transparent. The Commission, however, appears to exhibit little desire to (i) move beyond what it perceives to be its narrow statutory mandate and (ii) develop on its own initiative any new measures or indicators of unreliable delivery service.⁹ Consequently, although GAO and Congress may be interested in receiving reports that are consistent with a broader interpretation of reliability, they may have to wait until the Commission is given a more explicit statutory mandate for investigations that center on *reliability of delivery service provided by the Postal Service* and go beyond mere reliability of data.

II. Comments of PostCom/MMA

A. Postcom/MMA Clearly Regard the Statutory Mandate for Reliability as Entailing Something More than Just Accurate, Comprehensive Data

Recording usable data with respect to delivery of mail involves myriad details such as start-the-clock, stop-the-clock, critical entry times, etc. The detailed task of collecting accurate performance data can make it easy to adopt a narrow focus that, figuratively speaking, loses sight of the forest because of the trees. The following statement in PostCom/MMA's Comments thus provides the Commission with important perspective:

the main purposes of the USPS' service performance measurement and reporting system should be **to drive service improvement** and ensure that Market Dominant mail categories receive the level of service paid for in the price of the mail. Further, consistent, reliable, and predictable delivery service is imperative in order for mail to be an integral part of **multi-channel activities**. [*Id.* at 1 (emphasis added).]

⁹ See GAO Report, Appendix III: Comments from the Postal Regulatory Commission. *Op. cit.*, fn. 1, *supra*.

Postcom/MMA state that “*consistent, reliable, and predictable* delivery service is imperative in order for [advertising] mail to be an integral part of multi-channel activities.”

Id. Most through-the-mail advertisers can control the date when they enter their mail.

Consequently, for advertisers consistency, reliability, and predictability (all three of which relate chiefly to reliability of delivery service) are at least as important as timeliness, as Postcom/MMA point out. This is probably why the statute cites “reliability” as a separate consideration, distinct from “speed of delivery.” Yet the 2014 ACD makes no mention of reliability as a separate, distinct characteristic of performance. Instead, for Standard Mail the ACD seems to conflate reliability with timeliness of delivery.

In FY 2014 Standard Mail delivered on time ranged from 76.2 to 92.3 percent.¹⁰ The balance, 7.7 to 23.8 percent of the mail, was not delivered on-time. For this subset of mail, there is no indication of how unreliable delivery was in terms days late. Small wonder that PostCom/MMA want to “ensure that Market Dominant mail categories receive the level of service paid for in the price of the mail.” *Id.* Presumably, addressees sooner or later received the mail that was not delivered on-time. But after how many days? One? Three? Five or more? Small wonder that Postcom/MMA perceive problems making mail an integral part of multi-channel activities, especially when advertising campaigns released through competing channels can be timed to an exact day, perhaps even to the hour on some channels. So when will the reporting system begin to supply information on the extent of inconsistency and

¹⁰ *Id.*, see at 105, Table V-9, Service Performance Results for Standard Mail, FY2011 - FY2014.

unreliability of mail delivery? Not until reliability is viewed as a separate, measurable characteristic of delivery service, and something more than just quality of performance data.

Postcom/MMA recognize that simply collecting data, no matter how complete and accurate, and publishing those data in a series of tables will not, by themselves, drive the Postal Service to improve service performance. With the following statement, they clearly interpret the statutory requirement for reliability *broadly* as something more than just data quality, while implicitly rejecting a more *narrow view*.

We also recommend **additional analysis** be done for those geographical areas and facilities that are **consistently low-performing**. ... These **additional reports** would be given the PRC and the mailing industry to help better understand the data and what it means for on-time delivery moving forward. [*Id.* at 1 (emphasis added).]

Postcom/MMA's call for additional analyses recognizes that in order to deliver on the promise of more reliable delivery service, additional effort that goes beyond gathering and reporting data accurately is required. Their focus obviously is on having the Postal Service provide delivery service that is both (i) timely **and (ii) consistent**, with the performance measurement and reporting system viewed as a means to drive the Postal Service toward that most desirable operational end. Postcom/MMA is correct that expanding and improving the performance data base is important, but is not an end in itself. Rather, the fundamental goal should be to use the data base that underlies the performance measurement system to help drive operational improvements that will result in more consistent performance.

Implicit in the comments of Postcom/MMA is that:

- The first step in driving service improvement is to **identify** those operations where performance is below par and in need of improvement; and
- Existing tables of performance data **do not identify routinely** those below-par operations that may be in need of improvement.
- Some form of diagnostic analysis is needed to help complete and perfect the Postal Service's service performance measurement and reporting system.

What's missing, and is needed, are reports and analyses that ferret out and highlight undesired variation of delivery performance and help pinpoint where improvement is needed. One possible approach is illustrated by Figure 5 on page 25 of the GAO Report, as noted *supra*. Perhaps the Commission should consider requiring the Postal Service to submit tables or figures in which all postal districts are ranked in terms of performance, from worst to best (or vice versa), in a manner similar to the GAO Report. Or, alternatively, identify only the worst-performing districts. Or find still another way to achieve a result that identifies and ranks categories of mail and geographic areas where performance is the worst and deserving of attention by the Postal Service.¹¹

B. Existing Data May Be Adequate to Support Analysis of the Unreliability of Delivery Service for Many Categories of Mail

Since the purpose of recording and amassing performance data is to drive service improvements, then

- much of the *existing* performance data likely are sufficient to support investigations regarding many categories of mail whose

¹¹ See the Appendix for an illustration of reliability measures that could be developed and presented from existing performance data.

delivery is unreliable in varying degrees and whose reliability could stand improvement,

- accuracy and representativeness of most, even all, of the data supporting such conclusions most likely would be beyond challenge, but
- as currently designed, the reporting system fails to highlight the extent of unreliable delivery service given to any area or category of mail.

In any future dockets dealing with performance measurement, the Commission should consider giving more attention to development of alternative formats, analyses or other reports that would exploit and utilize the already extensive base of performance data in the most insightful and meaningful way possible. In that regard, even though existing data may be subject to further improvement (*see* Postcom/MMA's Comments at 5-15 addressing various technical matters), *an important asymmetry exists* as regards utilization of existing data.

Namely, suppose one drills into the existing data and finds no significant performance issues. Given the present incompleteness of the data — 45 percent are considered missing — one obviously cannot conclude that no performance issues exist elsewhere, which of course is why the Commission should continue its efforts to expand and improve the data base. But here is the asymmetry.

If one subjects the existing data to analysis and unearths categories of mail or geographic areas with significant operational issues such as highly erratic and inconsistent delivery — *e.g.*, delivery to rural areas, or delivery of Standard Mail — it is reasonable to conclude that problems exist which need to be addressed now — without waiting months or years for further expansion and quality improvements to the performance data base.

Expanding size of the data base by having more mail in measurement is unlikely, *per se*, to highlight or solve existing problems, or make them go away.

Finally, Postcom/MMA note that:

In addition to geographic area, there are areas of mail whose characteristics cause their service performance experience to differ from the rest of the measured mailstream. When only *aggregated data* is reported (such as data only on end-to-end and destination-entered mail), mailers and the PRC cannot evaluate whether these subsets of mail are meeting service performance standards. That is, their performance is rolled into the performance of the general mailstream, and the *aggregated data* obscures differences among the performance levels of particular areas of mail. [*Id.* at 4 (emphasis added).]

This is a good point. Most problems of unreliable service are local. It generally is understood that aggregation, especially when it is to the national level, can completely obscure virtually all local issues. It is a bit of a *non sequitur* for the ACD to say that after examining national data, and seeing no problems, Congress can rest assured that delivery service at the local level is satisfactory. The devil is in the details, as so often is the case.

/s/ John Haldi
John Haldi, Ph.D.
Haldi & Associates
118 E. 60th Street
Apartment 22C
New York, NY 10022-1107
(212) 935-0050
jhaldi4@gmail.com

Appendix

Development of Statistical Measures for Reliability of First-Class Single-Piece Letters FY 2015

The purpose of this appendix is to demonstrate how existing published data can be used to develop statistical measures that (i) give a profile of reliability, and (ii) lend themselves to targets against which actual performance and progress toward improving reliability can be measured. We start with the results and then explain their derivation. The mail category selected for illustrative purposes is First-Class single-piece letters, and the data are from FY 2015 — the same mail category and year in the GAO Report.

	(1) Mail Delivered On-time Target 95.0	(2) — Mail Not Delivered On-Time --- Target 5.0%	(3) Avg Days Late Target 1.0	(4) Standard Deviation Target 0.0
Mail Category				
Overnight	95.8%	4.2%	1.38	0.844
2-day	94.0%	6.0%	1.35	0.771
3-day	77.4%	22.7%	1.41	0.771

The data in columns 1 and 2 are the usual “speed of delivery” statistical measures, and the data in these columns are taken directly from the published performance table for First-Class single-piece letters. For mail that is delivered on-time, there are no reliability issues. The statistical measures in columns 3 and 4 focus on mail *not delivered on-time* and give a profile, or some insight, into the extent of reliable delivery service received by late-delivered First-Class single-piece letters in FY 2015, *i.e.*, all data here are from FY 2015.

Documenting the erratic delivery performance of Standard Mail in the manner shown here could be highly informative to concerned parties such as Postcom/MMA. To compute these statistics, extra effort will be required, as Postcom/MMA observed.

Targets. For FY 2015 the Postal Service established 95 percent as the target for on-time delivery of all First-Class Mail. This target means that no more than 5 percent should fail to be delivered on-time (col. 2).

The Postal Service has not published a target for when First-Class Mail not delivered on-time actually should be delivered. So, for purposes of this hypothetical exercise, an “ideal” target of 1.0 day(s) has been selected. This 1.0 day target means that any mail not delivered on-time should be delivered on the next delivery day. If all mail not delivered on-time were to be delivered exactly one day late, then the standard deviation of average days late would, under this “ideal” condition, be zero, *i.e.*, within the subset of late-delivered First-Class Single-Piece Letters there would be no deviation if late-delivered mail always arrives on the next delivery day.

In this hypothetical exercise, the ideal has been used for the target. If the above statistics on reliability were available for several years, as well as for other categories of mail, it might be appropriate to set more realistic targets for reliability, rather than the “ideal” (just as the target for on-time delivery of 95 percent is less than an ideal 100 percent but, hopefully, realistically achievable). The point here is: **statistical targets for reliability, distinct from speed of delivery, can be established. Each year, actual performance can be compared with targets for (i) average days late and (ii) standard deviation. Further, multi-year comparisons (when multi-year data become available) can help determine whether delivery service is improving or deteriorating in terms of reliability.**

Average days late. The develop the average days late for each mail category, one first needs the actual volume of late mail in each category, which can be derived from the above percentages and the published data, as shown below.

Mail Category	Mail Delivered On-time	Mail Delivered Late	
Overnight	95.8%	4.2%	
2-day	94.0%	6.0%	
3-day	77.4%	22.7%	
Volumes			
Overnight	559,212	23,214	552,726
2-day	1,851,858	118,204	1,970,062
3-day	<u>1,739,038</u>	<u>510,688</u>	<u>2,249,726</u>
Total	4,120,408	652,106	4,772,514

To compute the average days late for the overnight category, the percent of mail that was 1, 2, 3, and more and 3 days late is given in the performance table. At this point one assumption is necessary. Fortunately, for First-Class single-piece letters it is a modest assumption. Namely, all of the small amount of mail not delivered within 3 days (0.3 percent) is assumed to be delivered on the fourth day.¹² Computation of the average days late for the overnight category is as follows:

Days Late	Percent	No. of Pieces	Pieces x Days Late
1	3.3	18,240	18,240
2	0.5	2,764	5,527
3	0.1	553	1,658
4	<u>0.3</u>	<u>1,658</u>	<u>6,633</u>
	4.2	23,214	32,058

¹² If the performance data base contains the actual number of days late for pieces more than 3 days late, a computerized version of course would use those data.

The weighted average of days late = $32,058/23,214 = 1.38$. Computing the average days late for the 2-day and 3- day categories follows a procedure similar to that shown here, *i.e.*, using the distribution of days late and the mean days late for each respective category.

Variance and Standard Deviation. Computing the variance and standard deviation of late-delivered mail around the mean of late-delivered overnight mail is a straightforward exercise.

Days Late	No. of Pieces	Squared Dev. from Mean
1	18,240	2,634
2	2,764	1,062
3	553	1,141
4	<u>1,658</u>	<u>11,382</u>
	23,214	16,529
16,529/23,214 =		0.712 = variance
		0.844 = standard deviation

As stated at the outset, the purpose here is to demonstrate that the published performance data enable computation of statistical measures of reliability, measures that lend themselves to explicit performance targets. Standard, readily available statistical measures (mean, standard deviation) have been used, but other statistical measures could be devised in their stead. The point here is that (1) reliability is not timeliness, and (2) statistical measures of the reliability of delivery performance, separate and distinct from timeliness measures, can be constructed and presented using data now available in the published performance tables. These reliability measures can be computed for each product in the First-Class, Standard and Periodicals Classes of mail. They would add to understanding, and they might help pinpoint areas where the Postal Service needs to focus efforts to improve reliability of delivery. Effecting such improvements would help mailers and the Postal Service.